Area Inequalities for Embedded Disks Spanning Unknotted Curves

Joel Hass[1], Jeffrey C. Lagarias[1] and William P. Thurston[2]

February 1, 2008

Abstract

We show that a smooth unknotted curve in $\mathbb{R}^3$ satisfies an isoperimetric inequality that bounds the area of an embedded disk spanning the curve in terms of two parameters: the length L of the curve and the thickness r (maximal radius of an embedded tubular neighborhood) of the curve. For fixed length, the expression giving the upper bound on the area grows exponentially in $1/r^2$. In the direction of lower bounds, we give a sequence of length one curves with $r \to 0$ for which the area of any spanning disk is bounded from below by a function that grows exponentially with $1/r$. In particular, given any constant A, there is a smooth, unknotted length one curve for which the area of a smallest embedded spanning disk is greater than A.

1 Introduction

The classical isoperimetric inequality gives an upper bound for the area of the disk bounded by a curve in $\mathbb{R}^2$. For a simple closed curve of length L in the plane that encloses area A, it states that

$$4\pi A \leq L^2, \tag{1}$$

[1]The first two authors carried out some of this work while visiting the Institute for Advanced Study. The first author was partially supported by NSF grant DMS-0072348, and by a grant to the Institute by AMIAS.

[2]Partially supported by NSF grant DMS-9704286.

with equality if and only if the curve is a circle. There are many extensions of the isoperimetric theorem to higher dimensions; some of these are discussed in Almgren [1], Burago and Zalgaller [7], and Osserman [25].

There are two natural extensions of the isoperimetric inequality to a simple closed C^2-curve γ in $\mathbb{R}^3$ of length L.
(1) There exists an immersed smooth disk in $\mathbb{R}^3$, with γ as boundary, having area A with $4\pi A \leq L^2$. This disk may self-intersect and may intersect the curve γ, as indeed it must if the curve is knotted. If the curve is not a circle, then there exists such an immersed disk for which strict inequality holds.
(2) There exists an orientable smooth surface embedded in $\mathbb{R}^3$, with no restriction on its genus, with γ as boundary, having area A, with $4\pi A \leq L^2$. If the curve is not a circle, then there exists such a surface for which strict inequality holds.

Result (1) traces back to a result of Andre Weil in 1926 concerning isoperimetric inequalities for surfaces of non-positive Gaussian curvature. See also Beckenbach and Rado [4]. Result (2) traces back to a 1930 argument in Blaschke [5, p. 247], see Osserman [25, p. 1202]. For completeness we indicate proofs of these well-known results in an Appendix (§6). We derive (2) directly from (1). The hypotheses in (1) and (2) can be weakened to require only that γ be rectifiable; however we assume C^2-curves for ease of comparison with later results.

This paper treats the situation when the two conditions above are imposed simultaneously. That is, we suppose the curve γ is unknotted, and therefore is the boundary of an embedded disk, and ask for upper bounds on the area of some embedded spanning disk. Note that an embedded spanning disk attaining the minimal area under these conditions need not exist, because the embeddedness condition may not be preserved under limits. However there is still an infimum for the area of all embedded spanning disks, which we refer to as the minimal area of a spanning disk.

We will show that under these conditions there is no upper bound on the minimal area given in terms of the length L of γ alone. To obtain an upper bound on the area of some spanning disk we must therefore impose more geometrical control on γ than just bounding its length. For this we consider an additional geometric quantity, the thickness of γ.

The *thickness* $r(\gamma)$ of a C^2-curve γ in $\mathbb{R}^3$ is the supremum of r such that a radius r normal tubular neighborhood of the curve is embedded. This notion can be used to define a knot invariant. The supremum of the ratio $r(\gamma)/L(\gamma)$ over all C^2-curves γ having a given knot type defines a knot invariant called

the *thickness* of a knot, studied in Buck and Simon [6] and Litherland et al. [22], where it is related to various knot energies. For a C^2-closed curve γ the quantity $L(\gamma)/r(\gamma)$ is called the *ropelength* of γ, and the infimum of the ropelength over all C^2-representatives of a knot type, called the minimal ropelength, is the inverse of the thickness. Various properties of ropelength, including the existence of a minimizer in the class $C^{1,1}$, have been investigated recently [8].

Our first two results concern area bounds in terms of the length L and thickness r of an unknotted closed curve γ. In §3 we show that for an unknotted curve there is an upper bound for the area of a spanning disk depending on these quantities.

Theorem 1.1. *For any embedded closed unknotted C^2-curve γ in $\mathbb{R}^3$ having length L and thickness r, there exists a smooth embedded disk of area A having γ as boundary with*

$$A \leq (C_0)^{(L/r)^2} L^2,$$

where $C_0 > 1$ is a constant independent of γ, L and r.

In the other direction, we show that an upper bound of this type must be at least exponential in L/r, and that there is no upper bound on the area depending solely on L.

Theorem 1.2. *There is a constant $C_1 > 1$ and a sequence of unknotted, C^∞-curves $\{\Delta_n\}_{n=1,\ldots,\infty}$ embedded in $\mathbb{R}^3$, each having length $L_n = 1$ and thickness $r_n > 1/n$, with the property that for large enough n, any embedded disk spanning Δ_n has area*

$$A \geq (C_1)^n \geq (C_1)^{L_n/r_n} L_n^2.$$

We include the constant $L_n = 1$ for comparision to Theorem 1.1. These exponential lower bounds differ from the exponential upper bound in Theorem 1.2 in having an exponent linear in L_n/r_n rather than quadratic. Such examples necessarily have $1/n \leq r_n \leq c/\sqrt{n}$ for a constant c as $n \to \infty$, by comparing the lower bound $(C_1)^n$ with the upper bound in Theorem 1.1.

We next show that a bound on the length together with an upper bound on the curvature of an unknotted curve is also not sufficient to give an upper bound for the area of some spanning disk. Without loss of generality we may

normalize by setting the length of the curve to one.

Theorem 1.3. *There is an infinite sequence $\{\Gamma_n\}_{n=1,\ldots,\infty}$ of unknotted C^∞-curves of length $L_n = 1$ embedded in $\mathbb{R}^3$, each with (pointwise) curvature everywhere bounded above by a constant K_0, independent of n, such that any embedded disk spanning Γ_n has area*

$$A \geq n.$$

The proofs of Theorems 1.2 and 1.3 are both given in §5. They are proved via a detour into piecewise-linear topology, and an investigation of similar questions for piecewise-linear curves, i.e. embedded polygons in $\mathbb{R}^3$. For piecewise linear curves, one can consider combinatorial analogues of isoperimetric inequalities, in which the "length" is replaced by the number of edges n in a polygon, and the "area" is the number of triangles in an embedded triangulated surface having the polygon as boundary, as in [17]. We establish the following PL analogue of Theorem 1.1 above.

Theorem 1.4. *There is a constant $C_2 > 1$ such that given any polygonal unknotted curve P in $\mathbb{R}^3$ containing n line segments, there is a triangulated PL disk embedded in $\mathbb{R}^3$ having P as boundary that contains at most $(C_2)^{n^2}$ triangles.*

This result is established as part of Theorem 2.1 in §2. To deduce Theorem 1.1 from it, we show that we can obtain such a triangulated surface in which the area of each triangle is bounded above by a constant.

Hass, Snoeyink and Thurston [19] established the following PL analogue of Theorems 1.2 and 1.3.

Theorem 1.5. *There is a constant $C_3 > 1$ and an infinite sequence of unknotted polygonal curves $\{K_n\}_{n=1,\ldots,\infty}$ having between n and $23n$ segments, such that any triangulated PL disk embedded in $\mathbb{R}^3$ having K_n as boundary contains at least $(C_3)^n$ triangles.*

This result is not quite sufficient to deduce Theorem 1.2. It is necessary to strengthen the construction in [19], and we carry this out in Theorem 4.1 in §4. Using this strengthened construction, in §5 we show that a large number

of triangles occurring in the triangulated PL disk have their areas bounded below by a fixed positive constant, and that such the resulting lower bound carries over to smooth embedded disks with the same or nearby boundaries.

In the direction of PL combinatorial bounds, Hass and Lagarias [17] showed that there is a (qualitative) combinatorial analogue of the isoperimetric inequality (1) above in $\mathbb{R}^3$.

The plan of the paper is as follows. In §2 we derive the PL upper bound Theorem 1.4. From it we obtain upper bounds on the area of a piecewise-smooth embedded disk spanning a polygonal unknot P with n edges. In §3 we deduce Theorem 1.1 from these results. In §4 we review some results of [19], and construct a family of unknotted polygonal curves $\{K_n\}_{n=1,\ldots,\infty}$ that do not have PL spanning disks having few triangles, and have additional useful properties. The families of curves $\{\Delta_n\}_{n=1,\ldots,\infty}$ and $\{\Gamma_n\}_{n=1,\ldots,\infty}$ used to prove Theorems 1.2 and 1.3 are constructed by a smoothing process from these curves. In §5 we relate the PL complexity to area, and obtain lower bounds for the area of any spanning disk of Γ_n, resp. Δ_n, that increase to infinity with n, and deduce Theorems 1.2 and 1.3.

We note that the proofs of the lower bound estimates Theorems 1.2 and 1.3 in §4 and §5 can be read independently of the proofs of the upper bound results Theorem 1.1 and 1.4 in §2 and §3.

2 Upper bound: Minimal complexity of a PL embedded spanning disk

Let P be a polygonal, unknotted curve of length L in $\mathbb{R}^3$ having n edges. In this section we deduce upper bounds on the number of triangles in some embedded PL spanning disk for P, and for the minimal area of such a disk.

Theorem 2.1. *There exists a constant $C_4 > 1$ such that whenever P is a closed, unknotted, polygonal curve in $\mathbb{R}^3$ having n edges and length L, the following hold:*
(1) There exists an embedded PL disk spanning P whose number of triangles is at most $\frac{1}{32}(C_4)^{n^2}$.
(2) Such an embedded PL disk can be chosen to lie inside a ball of radius $4L$, and its area then satisfies

$$A \leq (C_4)^{n^2} L^2.$$

Theorem 2.1 is proved by a modification of ideas used in Hass and Lagarias [16]. That paper obtains upper bounds on the number of elementary moves required to move from an unknotted polygonal curve in the one-skeleton of a triangulated 3-manifold having t tetrahedra, to a single triangle, by a sequence of moves which deform the curve across three surfaces, an annulus, a torus and a disk. Here we need to produce a spanning disk. The construction of [16] can be modified to produce such a disk, with the bound on the number of triangles in the disk having the same functional form C^t as the bounds obtained in [16, Theorem 1.2], but with a different constant $C > 1$. A second necessary change is that we must construct a triangulation of a 3-ball containing the polygon P in its one-skeleton, and we do not have the freedom to move this curve; our construction obtains a quadratic upper bound for the number of tetrahedra in the constructed 3-ball triangulation.

Theorem 2.2. *There exists a constant $C_5 > 1$ such that whenever M is a triangulated PL 3-manifold with boundary, containing t tetrahedra, and K is a closed polygonal curve contained in the 1-skeleton of the interior of M that is unknotted in M, then there is a PL triangulated disk in M that has a subdivision of K as boundary, and that contains at most $(C_5)^t$ triangles.*

Proof: Theorem 1.2 of [16] states that there is a constant c_2, independent of M, such that if K is an unknot embedded in the 1-skeleton of interior(M), then K can be isotoped to a single triangle using at most $2^{c_2 t}$ elementary moves. The proof of this result deforms the original knot K to a triangle with three sequences of elementary moves, made inside three triangulated surfaces. A peripheral torus is constructed by taking two barycentric subdivisions and taking the torus that is the boundary of a regular neighborhood R_K of K. An embedded disk is used to construct elementary moves taking an essential curve K_2 on ∂R_K to a single triangle in M. On ∂R_K there is a curve K_1 with prescribed number of segments that is a longitude, and is isotopic to K in R_K by an isotopy with an explicitly computed bound on the number of elementary moves. An embedded annulus is used to go from K to the longitude K_1. The 2-torus ∂R_K is then used to define a sequence of elementary moves from K_1 to the boundary of a normal disk K_2. These moves define a homotopy between two embedded curves on ∂R_K.

The disk and the annulus used to construct the elementary moves are each embedded in M. However the elementary moves that connect K_1 to K_2 in ∂R_K may go back and forth over the torus ∂R_K many times, and the annulus that is swept out on ∂R_K may not be embedded. This problem is

rectified by a modification of the construction given in [16] that perturbs this swept out annulus to an embedding.

In Theorem 1.2 of [16] it is shown that:

(1) A normal disk with a longitude K_2 as boundary contains at most $2^{4608t+6}$ triangles.

(2) There is an embedded PL annulus S in the solid torus R_K whose two boundary components are another longitude K_1 and K, and which consists of at most 2^{1858t} triangles.

(3) The number of elementary moves needed to deform K_1 to K_2 on the torus ∂R_K is at most $2^{10^7 t}$.

Let n denote the number of elementary moves taking K_1 to K_2, so that $n \leq 2^{10^7 t}$ and let T denote the PL surface ∂R_K. We retriangulate M so as to insert a product $T \times [0,1]$ in place of the triangulated torus T, with $T \times 0$ and $T \times 1$ triangulated identically to T. The interior of $T \times [0,1]$ is triangulated as follows: The torus $T \times t_i$ is triangulated identically to $T \times 0$, $i = 1, \ldots, n$. We triangulate the product $T \times [t_i, t_{i+1}]$, $i = 1, \ldots, n-1$ in two steps. Begin by dividing this region into triangular prisms, one for each triangle on $T \times t_i$. Then divide each prism into 14 tetrahedra, by first dividing each rectangular face of the prism into four triangles with a common vertex at the center of the rectangle, and then coning the 14 triangles on the boundary of the prism to a new vertex at the center of the prism. We now construct an embedded annulus in $T \times [0,1]$ whose boundary on $T \times 0$ is K_2 and whose boundary on $T \times 1$ is K_1. The intersection of the annulus with $T \times (t_i, t_{i+1})$ consists of the image of K_1 in T after $i-1$ elementary moves $\times (t_i, t_{i+1})$. On $T \times t_i$ the annulus consists of the image of K_1 in T after $i-1$ elementary moves plus the single triangle swept out by the ith elementary move if the ith move is of type 2. If the ith move is of type 1 then it consists of image of K_1 in T after $i-1$ elementary moves plus or minus a single vertex corresponding to the ith move. The image of K_1 on $T \times t_i$ after i elementary moves for any i consists of less than $2^{10^7 t} 2^{1858t} = 2^{10^7 t + 1858t}$ edges. Each vertical annular piece lying in $T \times (t_i, t_{i+1})$ can be triangulated with four triangles for each vertical rectangle above an edge in $T \times t_i$, and each horizontal piece lying in a torus $T \times t_i$ contains at most one triangle. There are at most $2^{10^7 t}$ such annuli in $T \times [0,1]$, one for each elementary move. So the total number of triangles required to triangulate the annulus in $T \times [0,1]$ is at most $(2^{10^7 t + 1858t + 2})(2^{10^7 t})$. We obtain a total of at most $(2^{10^7 t})(2^{10^7 t + 1858t + 2}) \leq 2^{10^8 t - 1}$ triangles in the annulus connecting K_2 on $T \times 0$ to K_1 on $T \times 1$.

Adding to this the contributions of the disk and annulus in (1) and (2)

above results in a spanning disk containing less than $2^{10^8 t}$ triangles. We take $C_5 = 2^{10^8}$. $\square$

To apply this result, we construct a triangulated convex polyhedron B in $\mathbb{R}^3$ which contains the polygonal curve P in its one-skeleton; B will be the PL 3-manifold in the result above.

Lemma 2.3. *Let P be an unknotted polygonal curve in $\mathbb{R}^3$ having n vertices and length L. Then there exists a triangulated convex polyhedron B (with interior vertices) in $\mathbb{R}^3$, which contains P strictly in its interior, in its 1-skeleton, and which has the following properties.*

(1) The manifold B is contained in a sphere of radius $4L$ in $\mathbb{R}^3$.

(2) The number of tetrahedra in B is less than $290n^2 + 290n + 116$.

Proof: The proof is similar in spirit to the construction of a triangulated 3-manifold given in [16, Theorem 8.1]. However the argument in [16] must be modified, because here the polygonal curve P is fixed, and its vertices must be in the triangulation, whereas in [16] only the projection is specified, and there is freedom to move around the vertices to simplify the triangulation.

After rescaling P by the homothety $\mathbf{x} \to \mathbf{x}/L$, we may assume its length is 1. In this case P lies in the interior of a ball of radius $1/2$ in $\mathbb{R}^3$, which can without loss of generality be assumed to be centered at the origin. By rotating the coordinate system if necessary, we can suppose that the projection of P in the z-direction onto the xy-plane is a regular projection, and then, as in Lemma 7.1 of [18], this projection has less than n^2 crossings. The graph associated to the projection has straight-line edges, and its vertices consist of the projections of vertices of the polygon γ together with all crossing points.

As a preliminary step, we construct an augmented version of this projection graph in the $z = 0$ plane, which will be a triangulated planar graph that we call $\mathcal{G}$. Let S' denote the convex hull of the vertices of the projection graph, and add extra edges as necessary to fill out the boundary of the convex hull. Note that all crossing vertices in the projection lie strictly in the interior of the convex hull. Add new vertices to the graph by adding on each edge of this graph two new interior "special" vertices, spaced one-third of the way along the edge from its two endpoints. The resulting graph has at most $5n^2 + 5n$ vertices. This graph is planar with convex faces, and we triangulate it as follows. We first enclose each crossing vertex in four triangles, by adding the four edges connecting pairs of interior "special" vertices nearest to it. We then triangulate all remaining (convex) polygonal faces of the graph in an arbitrary fashion, adding no new vertices. As a final step we add a border

to this graph consisting of three points arranged in an equilateral triangle in the $z = 0$ plane at distance 3 from the origin, together with the three edges of this triangle. Note that each edge gets no closer to the origin than $3\sqrt{3}/2$, and so remains outside the convex hull S' of the projection graph, which is contained in the ball of radius $1/2$ around the origin. Finally we triangulate the outer region thus added by connecting each vertex on the boundary of S' to appropriate vertices of the equilateral triangle.

The resulting triangulated graph $\mathcal{G}$ has at most $5n^2+5n+3$ vertices, and so there are exactly $10n^2 + 10n + 4$ triangles in its triangulation. Its convex hull is contained in the circle of radius 3 centered at the origin.

To begin construction of the triangulated polyhedron, we start with two vertically translated copies of this augmented graph, in the planes $z = 1$ and $z = -1$. The convex hull of this set is a triangular prism, which will be the polyhedron B. We proceed to triangulate B, i.e. to dissect it into tetrahedra sharing common faces, such that γ is in the one-skeleton. We will use the triangulation of the upper and lower faces on $z = 2$ and $z = -2$ arising from the triangulation of the augmented graph.

We proceed to construct a simplicial complex $\mathcal{H}$ inside B which will include P in its 1-skeleton, and which vertically projects to $\mathcal{G}$ on the plane $z = 0$; we will include $\mathcal{H}$ in the triangulation and not include $\mathcal{G}$. The simplicial complex $\mathcal{H}$ is mostly two-dimensional, consisting of triangles lying over each triangle of $\mathcal{G}$ not including a crossing vertex, but thickened to a three-dimensional tetrahedron over each triangle of $\mathcal{G}$ that does include a crossing vertex. To begin, we add the polygon P, and add as new vertices along each edge of P the points that vertically project to crossing vertices of the graph; there are then two vertices on P that project to each crossing vertex of the graph. To these upper and lower crossing vertices of P add a vertical edge connecting them, which is called the vertical edge corresponding to the crossing. Then add extra vertices on this configuration, two along each edge, that project to the "special" vertices of the graph $\mathcal{G}$. We add the three equilateral vertices of $\mathcal{G}$ lying in the plane $z = 0$, plus the edges connecting them to the vertices in $\mathcal{H}$ that project to the corresponding vertices in $\mathcal{G}$ that they connect to. Each vertex in $\mathcal{G}$ that is not a crossing vertex has a unique "lifted" vertex in $\mathcal{H}$ that vertically projects to it, while crossing vertices have two "lifted" vertices, an upper and a lower one. Each triangle in the triangulation $\mathcal{G}$ not containing a crossing vertex then lifts uniquely to a triangle in $\mathcal{H}$ having the corresponding "lifted" vertices. Corresponding to each crossing vertex of $\mathcal{G}$ we create four tetrahedra in the simplicial complex $\mathcal{H}$, which share

the vertical edge corresponding to the crossing, and completely surround it. Each of the four tetrahedra has as vertices the upper and lower crossing vertex of P plus two "special" vertices whose projection gives a triangle in the triangulation $\mathcal{G}$. We have now specified the simplicial complex $\mathcal{H}$. It has the property that each top-dimensional simplex in it projects vertically to a unique triangle in $\mathcal{G}$ and each such triangle is covered once. Furthermore $\mathcal{H}$ has an "upper surface" and a "lower surface" obtained by extracting from each tetrahedron its triangular upper and lower face having the same vertical projection.

We now add vertical edges connecting each vertex in the "upper surface" of $\mathcal{H}$ to the corresponding vertex of the graph $\mathcal{G}$ in the plane $z = 2$. We can match triangular faces in this copy of $\mathcal{G}$ with the triangular face in the "upper surface" and their convex hull is a convex triangular prism, with boundary edges the vertical edges we just added. Each such triangular prism may be subdivided into fourteen tetrahedra, by first subdividing each of its three quadrilateral vertical faces into four triangles by adding a vertex at its barycenter, and then coning all the triangular faces to the barycenter of the triangular prism. Triangulations of adjacent prisms are automatically compatible (adjoining triangular faces match); this triangulates the "upper" portion of B. The "lower" portion is cut up into compatible tetrahedra in an exactly similar fashion, adding vertical edges. In this triangulation of B, we have 28 tetrahedra corresponding to each triangle in $\mathcal{G}$ that does not include any crossing vertex, and 29 tetrahedra for those faces that do. Thus this triangulation uses at most $290n^2 + 290n + 116$ tetrahedra.

We have obtained a linear triangulation of a convex region B contained in the ball of radius 4 such that P lies in the one-skeleton of B, strictly in the interior of B. Upon rescaling by the homothety $\mathbf{x} \to L\mathbf{x}$ we obtain the desired triangulation, with B inside a sphere of radius $4L$. $\square$

Remark. We do not know whether the bound $O(n^2)$ in the triangulation above is the optimal order of magnitude. Avis and ElGindy [3] give examples of sets of n points in $\mathbb{R}^3$, for arbitrarily large n, such that any triangulation of the convex hull including these points as vertices required at least $\Omega(n^2)$ tetrahedra. However, they also showed that for n points in general position (no four in a plane) there exists a triangulation of the convex hull using $O(n)$ tetrahedra. We do not know whether such an improvement to $O(n)$ is possible in the case above for general position vertices, because we impose the stronger requirement that P be in the 1-skeleton of the triangulation.

See [10] for a possible approach.

Proof of Theorem 2.1: (1) Let B be the triangulated ball constructed in Lemma 2.3, that has at most $t = 290n^2 + 290n + 116$ tetrahedra. By hypothesis P is unknotted in $\mathbb{R}^3$, and it follows that it is unknotted in the ball B, since an unknotting disk can be moved into B by an isotopy. By Theorem 2.2 there is an embedded PL spanning disk in B having P as boundary and containing at most $(C_5)^t$ triangles. This gives an upper bound of $\frac{1}{32}(C_4)^{n^2}$ triangles, with $(C_4) = 32(C_5)^{696}$, on noting that $n \geq 3$. This yields (1).

(2) The embedded PL spanning disk constructed in Lemma 2.3 is contained in a ball B of radius $4L$. Each triangle in it therefore has area at most $32L^2$. Thus the total area is bounded by $\frac{1}{32}(C_4)^{n^2}(32L^2)$, which yields (2). $\square$

Remark: We have spanned the polygonal curve P with a PL disk. Such disks can be approximated on their interior by smooth disks of no greater area, so Theorem 2.1 implies the same upper bound for the area of a smooth spanning disk for P.

3 Upper bound: Minimal area of a smooth embedded spanning disk

In this section we consider closed unknotted C^2-curves γ embedded in $\mathbb{R}^3$. We obtain upper bounds for the minimal area of an embedded smooth disk having the curve as boundary as a function of the length L and the thickness r of the curve. In §5 we will show that there is no upper bound for the minimal area of an embedded disk having the curve as boundary that is purely a function of the length and the maximal curvature of such a curve.

Let γ be a rectifiable closed curve in $\mathbb{R}^3$ of length L, and let $\{\gamma(s) : 0 \leq s \leq L\}$ be a parameterization of the curve by arc length, starting from a fixed point $\gamma(0)$ on the curve. Recall that an *r-neighborhood* T_r for a curve γ consists of all points in $\mathbb{R}^3$ within Euclidean distance r of some point on the curve. An r-neighborhood T_r is *tubular* if it is foliated by disjoint two-dimensional closed Euclidean disks $D_s(r)$ of radius r isometrically embedded

in $\mathbb{R}^3$, whose center is the point $\gamma(s)$, for $0 \leq s < L$, and each disk is normal to the curve γ at the point $\gamma(s)$. If so, we call the disks $D_s(r)$ *fibers* of the r-tubular neighborhood. Every closed C^2-curve embedded in $\mathbb{R}^3$ has an r-tubular neighborhood for some positive r, see [20, Theorem 5.1, pp. 109-110].

Theorem 3.1. *Suppose that γ is an unknotted C^2-curve in $\mathbb{R}^3$ of length L that has an r-tubular neighborhood. Then there is an embedded smooth disk having boundary γ whose area A satisfies*

$$A < (C_6)^{(L/r)^2} L^2,$$

with C_6 is a constant independent of γ.

The idea of the proof is to approximate γ by an inscribed polygonal closed curve P' which remains entirely inside a thinner normal tubular neighborhood and has $n \leq 32L/r + 1$ segments. By Theorem 2.1 there exists a PL spanning disk for P' having area at most $(C_4)^{n^2} L^2$. This is at most $C^{(L/r)^2} L^2$, for a different constant C. We deform this PL spanning disk to a spanning disk for γ using a homeomorphism τ of $\mathbb{R}^3$ that moves P' to γ, is the identity outside the r-tubular neighborhood of γ, and is a piecewise differentiable Lipschitz map with Lipschitz constant everywhere bounded by 5. It follows that the area of the resulting spanning disk for γ is larger by at most a factor of 25 over that of the PL spanning disk for P'. The Lipschitz map we construct is the identity outside a neighborhood of γ; this neighborhood is a sort of necklace of beads around γ, in which each bead is a right circular cone, with cone point on the curve γ, rather then a round ball. We first establish some preliminary results concerning the deformations of a curve to an inscribed polygon.

Lemma 3.2. *Let $\alpha : [0, ar] \to \mathbb{R}^3$ be a smooth, properly embedded unit speed C^2-curve with curvature bounded above by $1/r$, with $0 < a \leq 1$. The tangent vectors $T(s)$ of α lie inside a right circular cone with cone point at $\alpha(0)$, axis on the line in direction $\alpha'(0)$, and with cone angle $\theta_a = 2\arcsin a/2$ subtended to its axis. The entire curve $\alpha(s)$ lies within this closed cone.*

Proof: Let $T(s)$ denote the unit tangent at $0 \leq s \leq ar$ and $k(s) = ||T'(s)||$ the magnitude of the curvature vector. Then $||T'(s)|| = k(s) \leq 1/r$, and

$$||T(s) - T(0)|| = || \int_0^s T'(t)dt|| \leq \int_0^{ar} ||T'(t)||dt \leq (ar)(1/r) = a.$$

It follows from the geometry of the isosceles triangle formed by $T(0), T(s)$, and $T(s) - T(0)$ that all tangent vectors to $\alpha(s)$ lie in a cone of angle $\theta_a = 2\arcsin a/2$ centered around the line through $T(0)$. In particular,

$$T(0) \cdot T(s) \geq \cos(2\arcsin a/2) > 0$$

over this whole range, since $0 < a \leq 1$. The curve α itself also lies inside the cone with axis in the direction $T(0)$ and based at $\alpha(0)$. For if the curve ever left this cone then it would also leave a cone with a slightly larger cone angle, and would exit the boundary of this slightly wider cone transversely. The tangent vector of the α at such a point would not lie within the original cone. $\square$

We next compare the distance of two points on a curve with specified tubular neighborhood as measured along the curve and as measured in $\mathbb{R}^3$.

Lemma 3.3. *Let $0 < a < 1/2$ and let $\alpha : [0, ar] \to \mathbb{R}^3$ be a smooth, properly embedded unit speed C^2-curve with curvature everywhere bounded by $1/r$. Let e denote the line segment in $\mathbb{R}^3$ connecting $\alpha(0)$ and $\alpha(ar)$. Then α is contained in the right circular cone with cone point $\alpha(0)$, axis e, and that subtends the angle $2\theta_a = 4\arcsin(a/2)$ to its axis. Moreover*

$$(\mathit{length}(e))^2 \geq \cos(2\theta_a)\ (\mathit{length}(\alpha))^2.$$

Proof: Lemma 3.2 shows that two tangent vectors on $\alpha(s)$ differ by an angle at most $2\theta_a = 4\arcsin(a/2)$, so that their inner product satisfies $T(s_1) \cdot T(s_2) \geq \cos(2\theta_a) > 0$, using $0 < a \leq 1/2$ in the last inequality. Let T denote the unit tangent vector pointing from $\alpha(0)$ to $\alpha(ar)$, so that

$$T = \frac{1}{\mathrm{length}(e)}(\alpha(ar) - \alpha(0)) = \frac{1}{\mathrm{length}(e)} \int_0^{ar} T(s)ds.$$

Then for any $0 \leq s' \leq ar$, the inner product

$$\begin{aligned} T \cdot T(s') &= \frac{1}{\mathrm{length}(e)} \int_0^{ar} T(s) \cdot T(s')ds \\ &\geq \frac{ar}{\mathrm{length}(e)} \cos(2\theta_a) \\ &\geq \cos(2\theta_a). \end{aligned}$$

It follows that the entire curve lies in the right circular cone with vertex at $\alpha(0)$, axis the line segment e pointing in direction T, subtending angle

$2\theta_a = 4\arcsin(a/2)$, and with its endpoint $\alpha(ar)$ being the center of the base of the cone.

Now since length$(\alpha) = ar$ we obtain

$$\begin{aligned} \text{length}(e) &= ||\alpha(ar) - \alpha(0)|| = |\int_0^{ar} T(s) \cdot T(s')ds| \\ &= |\int_0^{ar} T(s') \cdot \left(\frac{1}{\text{length}(e)} \int_0^{ar} T(s)ds \right) ds'| \\ &\geq \frac{1}{\text{length}(e)} \int_0^{ar} \int_0^{ar} |T(s') \cdot T(s)|dsds' \\ &\geq \frac{(\text{length}(\alpha))^2}{\text{length}(e)} \cos(2\theta_a). \end{aligned}$$

This yields

$$(\text{length}(e))^2 \geq \cos(2\theta_a)(\text{length}(\alpha))^2,$$

as asserted. $\square$

In what follows we let γ denote a C^2-curve having a normal tubular neighborhood of radius r. The existence of the tubular neighborhood for γ implies that the absolute value of the curvature at each point of γ is at most $1/r$. (The points where the normal exponential map develops singularities are called the focal set. They occur at distance r for a curve at a point of (absolute) curvature $1/r$, see [11, p. 232].) Therefore Lemmas 3.2 and 3.3 are applicable.

We construct a particular polygonal curve P' approximating the C^2-curve γ as follows. Let P' in $\mathbb{R}^3$ have vertices on the curve γ taking $n = \lfloor \frac{32L}{r} \rfloor + 1$ equally spaced points $z_0, z_1, \ldots z_{n-1}$ along γ (measured by arclength of γ) and inscribing line segments connecting successive points. Let γ_j be the arc of γ between z_j and z_{j+1} and let $e_j = [z_j, z_{j+1}]$ be the line segment of P' between these points in $\mathbb{R}^3$, with last segment $e_{n-1} = [z_{n-1}, z_0]$. All arcs γ_j are of equal length L/n and we define the quantity a_j by $a_j r = \text{length}(\gamma_j) = L/n$, which implies that $1/33 \leq a_j \leq 1/32$. We will verify below that P' is embedded in $\mathbb{R}^3$.

Let C_j denote the right circular cone with cone point z_j, axis the line segment e_j, subtending the angle $4\theta_{a_j} = 8\arcsin(a_j/2)$, Note that C_j has base diameter no larger than $2\sin(4\theta_{a_j})(a_j r)$ since length$(e_j) \leq a_j r$, and the center of its base is z_{j+1}.

Lemma 3.4. *Let γ be a closed C^2-curve with tubular neighborhood of radius r. In the construction above, the arc γ_j lies inside the closed cone C_j. Distinct cones C_j and C_k with $j < k$ are disjoint unless $j = k+1$ or $j = n-1, k = 0$, and in this case the two cones intersect in the single point z_{j+1}, resp. $z = 0$ if $j = n-1$. The polygonal curve P' is embedded in $\mathbb{R}^3$.*

Proof: Lemma 3.3 applies to the arc γ_j to show that it is contained inside the cone C_j. The cone C_j subtends an angle to its axis of

$$4\theta_{a_j} = 8 \arcsin a_j/2 \le 8 \arcsin(1/64) < 0.14,$$

and the base of this cone has diameter at most

$$2\sin(4\theta_{a_j})(a_j r) \le 2(8\arcsin(1/64))r/32 < 0.01r.$$

If two cones C_j and C_k intersect, then they necessarily contain points $v_j \in \gamma_j$ and $v_k \in \gamma_k$ separated by distance at most $(0.02)r$. Let e be the line segment between v_j and v_k.

We claim that since γ has an r-tubular neighborhood, the shorter of the two arcs of γ connecting v_j and v_k has length at most $(0.03)r$. To see this, consider the arc of γ starting from $v_j = \gamma(s_0)$ to a variable point $w = \gamma(s')$ where s' is an arclength parametrization of γ that increases from $s = s_0$ as $\gamma(s')$ moves along the curve in one direction. The distance $||v_j - \gamma(s')||$ is initially increasing. Let $\gamma(s_2)$ be the first maximum point reached. At that point the tangent vector $T(s_2)$ is orthogonal to the line $[v_j, w]$, and part of this line lies in the normal disk of radius r around w so we must have $||v_j - w|| \ge r$. Similarly allowing s' to decrease from s we find that $||\gamma(s') - v_j||$ increases monotonically to the first point s_1 where $||v_j - \gamma(s_1)||$ reaches a maximum, and this maximum is at least r by the same reasoning.

Now the remaining part of the curve falls outside the normal tubular neighborhood of radius r determined by the arc $\{\gamma(s) : s_1 \le s \le s_2\}$. This normal tubular neighborhood includes all points within distance $r/4$ of v_j, for if v is a point with $||v - v_j|| \le r/4$ then the closest point $w = \gamma(s^*)$ on the arc $\{\gamma(s) : s_1 \le s \le s_2\}$ to v has $||w - v|| \le r/4$ from it, and attains a local minimum for the distance function from v. It is at distance at least $r/2 > r/4$ from each of the endpoints $\gamma(s_1), \gamma(s_2)$, since otherwise

$$||v_j - \gamma(s_1)|| \le ||v_j - w|| + ||w - v|| + ||v - \gamma(s_1)|| < r/4 + r/4 + r/2 = r,$$

a contradiction. The local minimality implies that $v - w$ is perpendicular to the tangent vector $T(s^*)$, so lies in the disk of radius r at $w = \gamma(s^*)$ normal

to $T(s^*)$. We conclude that all points on the curve γ within distance $r/2$ of v_j must lie in that part of the curve where $||v_j - \gamma(s')||$ is monotone increasing as $|s' - s_0|$ increases. The remark above shows that this holds at least for $|s' - s_0| \leq r$. Since $||v_j - v_k|| < 0.02r$, the point v_k must fall on this part of the curve. Now we can apply Lemma 3.3 to observe that on taking $a = 0.03$ the segment $e = [\gamma(0), \gamma(0.03r)]$ already satisfies

$$(\text{length}(e))^2 \geq \cos(2\theta_a)(ar)^2 \geq \cos(4 \arcsin 0.015)(0.03r)^2 > (0.02)^2 r^2.$$

We conclude that the shorter of the two arcs of γ connecting v_j and v_k has length at most $(0.03)r$, as asserted.

Since the length of each segment γ_j is at least $r/33 > 0.03r$, it follows that C_j and C_k are adjacent cones, i.e. $k = j + 1$ or $j = n - 1, k = 0$. Adjacent cones however intersect in a single point, the center of the base of one and the cone point of the other.

Finally, since each segment e_j of P' lies in a different cone, and overlaps are only possible at endpoints of the $e_j's$, the closed polygonal curve P' is embedded in $\mathbb{R}^3$. $\square$

The next two lemmas will be used to construct a Lipschitz homeomorphism inside the tubular neighborhood that carries P' to α. We start with a two-dimensional map. Given a two-dimensional Euclidean disk $D(d) = D(0, d)$ of radius d centered at $(0, 0)$ and a point $q = (x, y)$ in the disk, with $||q|| < d$, let $\sigma_{q,d} : D(d) \to D(d)$ be the map that sends the origin to q and that maps each line segment from the origin to a boundary point of $D(d)$ linearly to the line segment from q to the same boundary point. It is explicitly given by

$$\sigma_{q,d}(w) = w + (1 - ||w||/d)q.$$

Lemma 3.5. *The map $\sigma_{q,d} : D(d) \to D(d)$ is a Lipschitz homeomorphism with Lipschitz constant at most $1 + ||q||/d \leq 2$. Its inverse map $\sigma_{q,d}^{-1}$ is Lipschitz with Lipschitz constant at most $\dfrac{1}{1 - (||q||/d)}$. If q_1 and q_2 lie in the disk $D(d)$, then for any point $w \in D(d)$*

$$||\sigma_{q_1,d}(w) - \sigma_{q_2,d}(w)|| \leq ||q_1 - q_2||.$$

Proof: The map $\sigma_{q,d}$ is clearly a homeomorphism, which leaves the boundary of $D(d)$ fixed. It remains to estimate a Lipschitz bound. For $w \in D(d)$ we have

$$\sigma_{q,d}(w) = w + (1 - ||q||/d)q.$$

We calculate that

$$\begin{aligned}||\sigma_{q,d}(w_2) - \sigma_{q,d}(w_1)|| &\leq ||w_1 - w_2|| + \frac{||q||}{d}(||w_1|| - ||w_2||) \\ &\leq ||w_1 - w_2||(1 + ||q||/d).\end{aligned}$$

Similarly $||\sigma_{q,d}(w_2)-\sigma_{q,d}(w_1)|| \geq ||w_1-w_2||(1-||q||/d)$. So $\sigma_{q,d}$ has Lipschitz constant bounded above by $1+||q||/d$ and below by $1-||q||/d$. It follows that the inverse map $\sigma_{q,d}^{-1}$ has Lipschitz constant bounded above by $\dfrac{1}{1-(||q||/d)}$.

Finally, we have

$$||\sigma_{q_1,d}(w) - \sigma_{q_2,d}(w)|| = |(1 - ||w||/d)|||q_1 - q_2|| \leq ||q_1 - q_2||,$$

as asserted. $\square$

Next we give a parametrized generalization of the map of Lemma 3.5 which is a three-dimensional map. We define in (x, y, t)-space the right circular cone

$$C(b, h) := \{(x, y, t) : x^2 + y^2 \leq b^2t^2, \quad 0 \leq t \leq h\}$$

of height h and base of radius bh, centered on the t-axis, and suppose that $0 < b \leq 1$. We consider a curve $\alpha(t) = \{(x(t), y(t), t) : 0 \leq t \leq h\}$ parametrized by the t-variable which is strictly monotone increasing in t, with $\alpha(0) = (0, 0, 0)$ and $\alpha(h) = (0, 0, h)$. and we assume that the whole curve is contained in the right circular cone $C(b, h)$. Concerning the smoothness of this curve we assume only that $t \to \alpha(t)$ is Lipschitz-continuous with Lipschitz constant $C_L(\alpha)$, i.e. that

$$||\alpha(t_1) - \alpha(t_2)|| \leq C_L(\alpha)|t_1 - t_2|$$

holds for $0 \leq t_1, t_2 \leq h$.

Now, associated to this curve α, we define a homeomorphism $\tau = \tau_\alpha$ of $\mathbb{R}^3$ supported in the cone $C(b, h)$, which leaves its boundary fixed. It takes horizontal disks with fixed z-coordinate into themselves moving each of them by the map $\sigma_{\alpha(t),bt}$ of Lemma 3.5, and moves the point $(0, 0, t)$ on the axis of the cone to $\alpha(t) = (x(t), y(t), t)$. This homeomorphism is defined inside the cone $C(b, h)$ by

$$\tau(x, y, t) := \sigma_{(x(t),y(t)),h(t)}(x, y, t), \qquad \text{with} \quad h(t) = bt,$$

and is the identity map on the boundary of the cone.

Lemma 3.6. *Suppose $0 < b \le 1$. The map $\tau : C(b,h) \to C(b,h)$ associated to the curve $\alpha(t)$ is a homeomorphism of $C(b,h)$, leaving its boundary fixed. This map is a Lipschitz map with Lipschitz constant $C_L(\tau)$ satisfying*

$$C_L(\tau) \le 4\sqrt{1+b^2} + C_L(\alpha) + 2.$$

Proof: Let $u = (x_0, y_0, t_0), v = (x_1, y_1, t_1)$ be two points in $C(b,h)$. If they lie at the same height $t_0 = t_1$, then since $\alpha(t_0)$ lies inside C_L, Lemma 3.5 yields the Lipschitz bound $||\tau(u) - \tau(v)|| \le 2||u - v||$.

Now if $t_0 \ne t_1$ we may arrange that $t^* := t_1 - t_0 > 0$, by exchanging u and v if necessary. Set $B := \sqrt{1+b^2}$, and our object is to show that

$$||\tau(u) - \tau(v)|| \le (4B + C_L(\alpha) + 2)||u - v||.$$

We study the effect of radial projection using the cone vertex $(0,0,0)$. Given a point $w = (x, y, t_0)$ in the plane $z = t_0$ we let $P(w) = (t_1/t_0)w$ be its radial projection to the plane $t = t_1$. Then for all such points w in the cone $C(g,h)$ we have

$$t^* \le ||P(w) - w|| \le (\sqrt{1+b^2})t^* = Bt^*,$$

the upper bound being tight for w on the edge of the cone.

Since $P(u)$ and v both lie in the cone in the same plane $t = t_1$, we have

$$\begin{aligned} ||\tau(u) - \tau(v)|| &\le ||\tau(u) - P(\tau(u))|| + ||P(\tau(u)) - \tau(P(u))|| + ||\tau(P(u)) - \tau(v)|| \\ &\le Bt^* + ||P(\tau(u)) - \tau(P(u))|| + 2||P(u) - v||. \end{aligned}$$

To bound the term $||P(\tau(u)) - \tau(P(u))||$ we observe that in terms of the function of Lemma 3.5 applied in the plane $t = t_1$ of the cone $C(b,h)$ we have $P(\tau(u)) = \sigma_{P(\alpha(t_0)),bt_1}(P(u))$ while $\tau(P(u)) = \sigma_{\alpha(t_1),bt_1}(P(u))$. Then

$$\begin{aligned} ||\sigma_{P(\alpha(t_0)),bt_1}(P(u)) - \sigma_{\alpha(t_1),bt_1}(P(u))|| &\le ||P(\alpha(t_0)) - \alpha(t_1)|| \\ &\le ||P(\alpha(t_0)) - \alpha(t_0)|| + ||\alpha(t_0) - \alpha(t_1)|| \\ &\le Bt^* + C_L(\alpha)t^*, \end{aligned}$$

where the first line above follows using Lemma 3.5. Substituting this in the previous inequality yields

$$||\tau(u) - \tau(v)|| \le (2B + C_L(\alpha))\, t^* + 2||P(u) - v||.$$

On the other hand we have $||u - v|| \geq t$ and

$$||u - v|| \geq ||P(u) - v|| - ||u - P(u)|| \geq ||P(u) - v|| - Bt^*.$$

We average these two inequalities in a suitable ratio to obtain

$$\begin{aligned}
||u - v|| &\geq \frac{4B + C_L(\alpha)}{4B + C_L(\alpha) + 2} t^* + \frac{2}{4B + C_L(\alpha) + 2}(||P(u) - v|| - Bt^*) \\
&\geq \frac{1}{4B + C_L(\alpha) + 2}\left((2B + C_L(\alpha))t^* + 2||P(u) - v||\right) \\
&\geq \frac{1}{4B + C_L(\alpha) + 2}||\tau(u) - \tau(v)||,
\end{aligned}$$

which gives the Lipschitz bound for τ. $\square$

Proof of Theorem 3.1: The curve γ has a normal tubular neighborhood of radius r. As described above we inscribe in γ a closed polygon P' with n edges, with $n \leq \frac{32L}{r}$, labelling the points $z_j = \gamma(s_j)$, the j-th edge $e_j = [z_j, z_{j+1}]$ of P', and γ_j the corresponding arc of the curve between $\gamma(s_j)$ and $\gamma(s_{j+1})$, as described above. We also have the set of n right circular cones $\{C_j : 0 \leq j \leq n-1\}$, where C_j has the segment e_j as its axis. Now Lemma 3.4 applies to show that P' is embedded in $\mathbb{R}^3$ and that these cones are disjoint except at the points z_j.

We define a homeomorphism $\tau : \mathbb{R}^3 \to \mathbb{R}^3$, which is the identity outside all the cones C_j and which inside each cone C_j is given by the homeomorphism τ_j of C_j given by Lemma 3.6, taking the arc $\alpha(t)$ to be the arc γ_j, reparameterized using the variable t that measures distance along e_j, rather than the arc length variable s. With this reparameterization the map $t \to \alpha(t)$ is a Lipschitz map. To bound the Lipschitz constant we use results from the proof of Lemma 3.4. It shows that the tangent vector T in the direction of e_j satisfies $1 \geq T \cdot T(s') \geq \cos\theta_a$ for all tangent vectors on the arc γ_j, where $0 < \theta_a \leq 8 \arcsin 1/64$ which yields $\cos\theta_a > 0.9$. This implies that the Lipschitz constant $C_L(\alpha) \leq 2$.

We conclude using Lemma 3.6 that τ is a Lipschitz map and obtain a bound for its Lipschitz constant inside the individual cones C_j. We take $h = \text{length}(e)$ and $b \leq \sin(8\arcsin(\theta_{a_j}/2)) \leq \sin(0.14) < 0.14$, which yields the bound $C_L(\tau) \leq 4(0.14) + 2 + 2 \leq 5$. Since τ is the identity map outside the cones, this Lipschitz bound carries over to $\tau : \mathbb{R}^3 \to \mathbb{R}^3$, using the triangle inequality.

Now consider the triangulated PL surface Σ' having P' as boundary given by Theorem 2.1, having area A' bounded by

$$A' < (C_4)^{n^2}(L')^2,$$

where L' is the length of P'. Using the bound $n \le 33L/r$ and that $L' \le L$ and $L/r > 1$, this gives

$$A' < (C_4)^{(33L/r)^2}(L')^2 \le (C_0)^{(L/r)^2}L^2,$$

for an enlarged constant $C_0 = (C_4)^{33}$.

We can assume, possibly after a small perturbation that induces an arbitrarily small increase in the area of Σ', that Σ' intersects the boundary of the $r/2$-tubular neighborhood of γ and the boundary of each ball B_j transversely. This follows from the fact that the space of smooth disks is dense in the space of piecewise smooth disks in $\mathbb{R}^3$, and the area functional is continuous on the space of piecewise smooth disks with the smooth topology. See for example Kosinski [21, Theorem 2.5].

Applying the homeomorphism τ then gives a piecewise-C^2 embedded surface $\Sigma := \tau(\Sigma')$ that has γ as boundary and whose area is increased from the area A of Σ' by at most a factor of 25, using the bound of 5 on the Lipschitz constant of the map τ. (A Lipschitz map $\tau : \mathbb{R}^n \to \mathbb{R}^n$ with Lipschitz constant $C_L(\tau)$ can stretch the volume of an m-dimensional submanifold by at most a factor $C_L(\tau)^m$, see Federer[13, 2.10.11], Morgan [23, Prop. 3.5].) The fact that Σ is piecewise C^2 follows using the fact that the map τ is piecewise C^2, with its smoothness inherited from the smoothness of the curve $\gamma(s)$. Thus, enlarging the constant C_0 to a suitable C_6 we have

$$A < \frac{1}{2}C_6^{(L/r)^2}L^2.$$

Finally, we can perturb the interior of Σ to obtain a smooth surface while again changing its area by an arbitrarily small amount, and without moving the boundary γ, see [21, Theorem 2.5]. This yields the theorem, where we give up another factor of 2 on the right side of the bound above to absorb the effect of the perturbation. $\square$

4 Lower bound: Large complexity minimal PL spanning disks

In this section we construct a series of polygonal unknotted curves K_n in $\mathbb{R}^3$ whose minimal spanning disks must have exponential complexity, according to various measures. This is an extension of the results of [19].

Theorem 4.1. *There exists a sequence of unknotted polygonal curves $\{K_n\}$ in $\mathbb{R}^3$, for $n \geq 1$, having the following properties.*
(1) K_n contains at most $24n + 40$ edges.
(2) Any embedded, piecewise-smooth disk spanning K_n intersects the y-axis in at least 2^{n-1} points.
(3) Any embedded PL triangulated disk D_n bounded by K_n contains at least 2^{n-1} triangles.
(4) Each K_n has length one and there exists a constant $r_0 > 0$ independent of n such that each K_n is disjoint from the solid cylinder $Cyl(r_0)$ of radius r_0 around the y-axis.

Three curves in this sequence are shown in Figure 1.

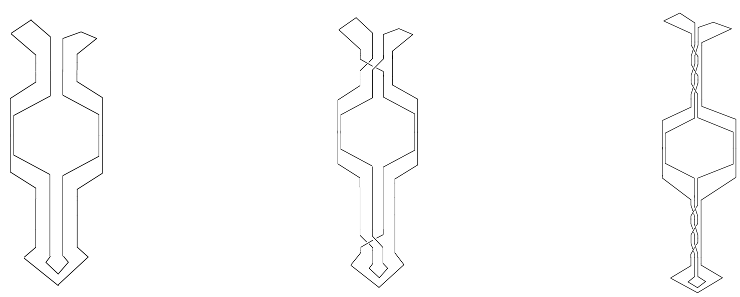

Figure 1: The curves K_0, K_1 and K_3.

Remark. The construction of the curves K_n in Theorem 4.1 differs from that in [19], in order to make condition (4) hold. The main focus of [19] was to construct examples of curves with $O(n)$ edges satisfying property (3), while in the application here it is properties (2) and (4) that are needed.
Proof: The assertions (1)-(3) of the theorem are invariant under homothety $\mathbf{x} \to \lambda\mathbf{x}$, so restricting to curves K_n of length one results in no loss of generality.

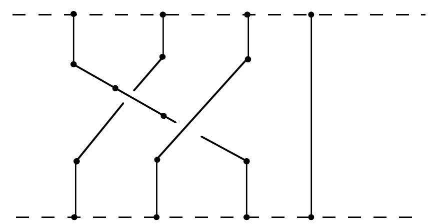

Figure 2: The braid $\alpha = \sigma_1\sigma_2^{-1}$ can be embedded in $\mathbb{R}^3$ using 12 segments so that the top and bottom ends of each arc are vertical segments on the xz-plane. The leftmost strand is made from five segments, and rises above and below the xz-plane. The other three strands lie completely in the xz-plane.

The construction of K_n begins with the 4-braid $\alpha = \sigma_1\sigma_2^{-1}$ shown in Figure 2. We construct α to consist of vertical segments near its endpoints and to have (total) length one. We stack n copies of α to get a braid α^n, and then rescale by a homothety to get a braid $\alpha_n = \dfrac{1}{3n}\alpha^n$ whose length is 1/3. Similarly we stack n copies of α^{-1} and scale to get a braid α_{-n} of length 1/3. The curve K_n is constructed by starting with a copy of each of α_n, α_{-n} and adding eight arcs to obtain a closed curve in $\mathbb{R}^3$. The eight arcs are shown in Figure 3.

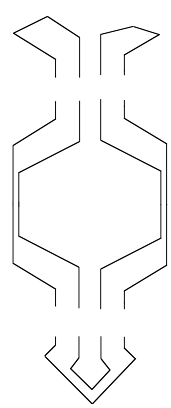

Figure 3: Arcs $a_n, b_n, \ldots, h_n$ used in forming the top, middle and bottom pieces of K_n. As n increases, the horizontal separation between the endpoints of the arcs decreases to 0, but their total length remains equal to 1/3. All arcs lie in the xz-plane and are disjoint from a disk in this plane of radius r_0 centered at the origin.

The added eight arcs all lie in the xz-plane, as shown in Figure 3. They consist of two arcs a_n, b_n connecting the first and second and the third and fourth endpoints at the top of α_n, four arcs c_n, d_n, e_n, f_n connecting the bottom four endpoints of α_n to the top four endpoints of α_{-n}, and two arcs g_n, h_n connecting the first and fourth and the second and third endpoints of the bottom of α_{-n}. Forty line segments are used to construct $a_n, b_n, c_n, d_n, e_n, f_n, g_n, h_n$. The resulting closed curves K_n, shown in Figure 1, are each unknotted, since the braid $\alpha^n\alpha^{-n}$ is trivial, and each K_n uses at most $24n + 40$ segments, with a contribution of $12n$ coming from each of α_n, α_{-n}. For each n, we construct the eight added arcs so that they have total length $1/3$, and so that they are disjoint from a solid cylinder of fixed radius r_0. With the total length of all the segments set to one, a calculation shows that we can take r_0 to be $1/100$. See Figure 4.

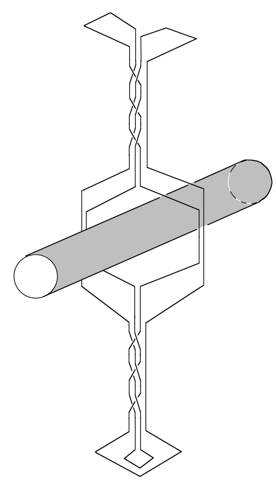

Figure 4: Each K_n is disjoint from a cylinder of radius r_0 around the y-axis.

As $n \to \infty$ the separation between the top four points of α_n converges to zero, and the endpoints of a_n and b_n which are identified to them also converge. Endpoints of $c_n, \ldots, h_n$ converge similarly. Each curve K_n is a length one curve disjoint from the cylinder $Cyl(r_0)$ of radius r_0 around the y-axis.

Now consider an arbitrary piecewise smooth disk D_n with boundary K_n. The intersections of D_n with the level sets $\{z = c\}$ can be analyzed using Morse Theory and the theory of train tracks and surface diffeomorphisms. This analysis is carried out in Theorem 1 of [19], where it is shown that the

y-axis must intersect D_n in at least 2^{n-1} points. All cases of Theorem 4.1 have now been established. $\square$

5 Lower Bound: Large minimal area embedded spanning disks

In this section we show that there is no upper bound on the minimal area of any embedded disk that depends only on the length L of the curve, as stated in Theorem 1.3. This is based on the sequence of curves K_n constructed in Theorem 4.1.

Lemma 5.1. *Let $\{K_n\}_{n=1,\ldots,\infty}$ be the sequence of length one curves constructed in Theorem 4.1. There is a constant $c_7 > 1$ such that any embedded disk D_n having boundary K_n has area A satisfying*

$$A > c_7 2^n.$$

The same lower bound holds for the area of any embedded disk having boundary a curve γ' that is isotopic to K_n by an isotopy that fixes the cylinder $Cyl(r_0)$ of radius r_0 around the y-axis.

Proof: Suppose a line L in $Cyl(r_0)$ parallel to its axis intersects D_n in less than 2^{n-1} points. There is an isotopy of $\mathbb{R}^3$ fixing K_n and carrying L to the y-axis. The image of D_n under this isotopy would be a disk D'_n spanning K_n and intersecting the y-axis in less than 2^{n-1} points, contradicting Theorem 1 of [19]. So any such line in the cylinder $Cyl(r_0)$ intersects D_n in at least 2^{n-1} points. The area of a perpendicular cross-section of $Cyl(r_0)$ is πr_0^2, so the projection of D_n to the xz-plane has area at least $2^{n-1}\pi r_0^2$, as does D_n itself. We choose $c_7 = \pi r_0^2/2$.

Now suppose that L_n is a curve isotopic to K_n in the complement of $Cyl(r_0)$ and E_n is a disk spanning L_n. Then there is an isotopy of $\mathbb{R}^3$ carrying E_n to a disk spanning K_n that restricts to the identity on $Cyl(r_0)$. The area of E_n in $Cyl(r_0)$ is at least $c_7 2^n$ by the previous argument. $\square$

To obtain Δ_n we will construct a smooth curve of length one that approximates K_n, and in particular remains disjoint from $Cyl(r_0)$ and has thickness greater than $1/n$, as shown in Figure 5.

Proof of Theorem 1.2: It suffices to find curves Δ_n having length at most one with the other claimed properties, as rescaling by a homothety to make these curves have length precisely one increases both the thickness and the area of any spanning disk. We will approximate each K_n by a smooth curve $\tilde{\Delta}_n$, isotopic to K_n in the complement of the cylinder $Cyl(r_0)$. Finally we will take $\Delta_n := \tilde{\Delta}_{[cn]+1}$ for an appropriate constant $c \geq 1$.

We first smooth the braids α and α^{-1} by replacing neighborhoods of their vertices by smooth arcs, getting a smooth braid β of thickness $k > 0$ and curvature bounded above by a constant K_0. We do this smoothing so that β still consists of vertical arcs near its endpoints. Then there is a constant $r_1 > 0$ with the property that the thickness of β is greater than r_1, and each vertical segment near an endpoint of β has length greater than r_1. Form β^n and β^{-n} as before by stacking copies of β or β^{-1}. Since copies of β are identified along points that have neighborhoods coinciding with vertical segments, the resulting curve is C^∞. Moreover the thickness of β^n and β^{-n} is still greater than r_1, since distinct copies of β have r_1-tubular neighborhoods with disjoint interiors. Scaling by a homothety gives $\beta_n = \dfrac{1}{3n}\beta^n$, with length $1/3$, and similarly β_{-n}. The thickness of each of β_n and β_{-n} is greater than $\dfrac{r_1}{3n}$. We can also smooth the curves $a_n, \ldots, h_n$ in neighborhoods of their internal vertices, to obtain nearby arcs $a_n', \ldots, h_n'$ that are smooth with arcs near the endpoints remaining as straight vertical segments, with the length of each straight segment no less than a constant r_2 (which does not depend on n). The approximating arcs are chosen sufficiently close to maintain their disjointness from $Cyl(r_0)$, and with their total length remaining less than $1/3$. Then combining β_n, β_{-n} with $a_n', \ldots, h_n'$ gives a closed curve $\tilde{\Delta}_n$ of length one that is isotopic to K_n in the complement of $Cyl(r_0)$, and the thickness of $\tilde{\Delta}_n$ is given by $\dfrac{k}{3n}$, where k is the smaller of r_1, r_2.

We now establish a lower bound on the area A of any embedded spanning disk for $\tilde{\Delta}_n$. Lemma 5.1 implies that there is a $C_0 > 1$ such that for large enough n, any embedded disk with boundary $\tilde{\Delta}_n$ has area A greater than ${C_0}^n$.

Note that $k < 1$ since $\tilde{\Delta}_n$ has length one and thickness less than k. Thus we can rescale n to eliminate the constant k appearing in the thickness bound by taking a subsequence $\Delta_n := \tilde{\Delta}_{[kn]+1}$, and setting $C_1 = (C_0)^k$. We then have minimal area $A > (C_1)^n$. $\square$

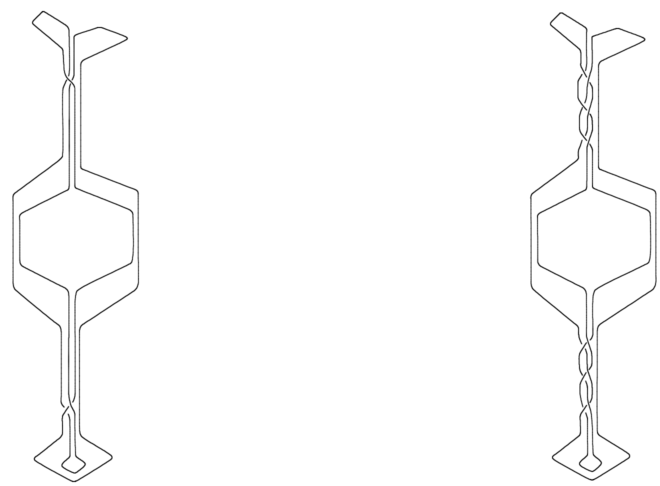

Figure 5: Each $\tilde{\Delta}_n$ is a smooth curve that has length one and thickness at least k/n. $\tilde{\Delta}_1$ and $\tilde{\Delta}_3$ are shown.

In the next result we show how to deform K_n to construct a similar sequence of smooth curves Γ_n having uniformly bounded curvature. Recall that the curvature function for a C^2-curve embedded in $\mathbb{R}^3$ is computed by taking an arclength parameterization $\gamma : [0, L] \to \mathbb{R}^3$, so that $||\frac{d}{dt}\gamma(t)|| = 1$, and the curvature function is then $\kappa(t) = ||\frac{d^2}{dt^2}\gamma(t)||$.

Proof of Theorem 1.3: The construction of the family $\{\Gamma_n\}_{n=1,\ldots,\infty}$ is similar to that of $\{\Delta_n\}_{n=1,\ldots,\infty}$. We again use smoothings $a'_n, \ldots, h'_n$ of the arcs $a_n, \ldots, h_n$ to curves that lie nearby them, and assert that the approximating curves can be chosen to have curvature at each point bounded above by a fixed constant K. To see this we use the fact that a sequence of smooth curves converging to a smooth limit curve have curvatures converging pointwise to the curvature of the limit curve. For any curve with monotonically increasing z-component, a deformation of the curve by a diffeomorphism $g_\lambda(x, y, z) := (\lambda x, \lambda y, z)$ carries it to a curve that, for λ small, lies very close to to the z-axis. For λ sufficiently small the curvature of the image curve under g_λ is uniformly close to zero. We choose $\lambda_n \to 0$ to be a sequence converging to 0 sufficiently fast so that the curvature of $\beta_n = g_{\lambda_n}(\beta^n) \to 0$, where β^n is as before, and similarly we define β_{-n}. We also smooth the curves $a_n, \ldots, h_n$ in neighborhoods of their internal vertices, to obtain nearby arcs $a'_n, \ldots, h'_n$

that are smooth with arcs near the endpoints remaining straight vertical segments. While these vertical segments all converge to the z-axis as $n \to \infty$, the curvature of each of $a'_n, \ldots, h'_n$ remains uniformly bounded. One way to see this explicitly is to choose smooth limiting curves $a'_\infty, \ldots, h'_\infty$ of bounded curvature that are not embedded, but rather have arcs near their endpoints that agree with the z-axis. Any sequence of smooth curves $a'_n, \ldots, h'_n$ that converge to these curves smoothly will have uniformly bounded curvature. Then the curvature of the curves Γ_n actually converges to 0 at points corresponding to the braids β^n, β_{-n}, and is uniformly bounded at other points. The curvature is zero along the straight segments where $a'_n, \ldots, h'_n$ join β^n and β_{-n}.

Since the required area for a spanning disk goes to infinity as $n \to \infty$, we can rechoose Γ_n to be a suitable subsequence to have the required property $A > n$. This concludes the proof of Theorem 1.3. $\square$

6 Appendix: Isoperimetric theorems for curves in $\mathbb{R}^3$

This appendix establishes two versions of the isoperimetric theorem in $\mathbb{R}^3$ as stated as (1) and (2) in §1. Result (1) follows from the solution of the Plateau problem, due to Douglas and Rado, [12], [26], and on a result of Carleman on isoperimetric inequalities for minimal disks [9]. See also Reid [27]. We deduce (2) from (1), using standard cut and paste methods of 3-manifold topology, which can be used to convert an immersed surface to an embedded surface, possibly with an increase in genus, but maintaining orientability. Such a cut and paste can be achieved without increasing area by a "rounding of creases". Alternate approaches to proving (2) are possible based on work of Blaschke [5, p. 247], which was the original approach to (2). The discussion in Osserman [25, p. 1202] indicates that Blaschke's argument is heuristic, but can be made valid for area-minimizing immersed surfaces having the curve as boundary.

Theorem 6.1. *Let γ be a simple closed C^2-curve in $\mathbb{R}^3$ of length L. Then there exists an immersed disk having area A in $\mathbb{R}^3$, with γ as boundary, such*

that

$$4\pi A \leq L^2.$$

If the curve is not a circle, then there exists such an immersed disk for which strict inequality holds.

Proof of Theorem 6.1: Span γ by a least area disk D, whose existence is guaranteed for rectifiable γ by the solution of the Plateau problem [12], [26]. The regularity results of Osserman [24] and Gulliver [15] show that D is a smooth immersion in its interior. (The interior of D may transversely intersect γ, as indeed it must if γ is knotted.) The argument in Reid [27] (see also [9]) shows that $4\pi A \leq L^2$, with equality if and only if γ is a circle. $\square$

Theorem 6.2. *Let γ be an embedded closed C^2-curve in $\mathbb{R}^3$ of length L. Then there exists an embedded orientable smooth spanning surface for γ in $\mathbb{R}^3$ having area A, with*

$$4\pi A \leq L^2. \tag{2}$$

If the curve is not a circle, then there exists such a surface for which strict inequality holds.

Proof of Theorem 6.2: Apply Theorem 6.1 to get an immersed disk D spanning γ satisfying the isoperimetric inequality (1). If D is not embedded then perturb it slightly so that its self-intersections are in general position, i.e. transverse with finitely many triple points. Since smooth general position maps are dense among C^2 maps, this can be done with an arbitrarily small increase in area. The self-intersections then consist of a finite number of immersed double curves. There is a uniquely defined surface F obtained by oriented cut-and-paste along these curves. An arbitrarily small perturbation of F moves it to an embedded surface. The resulting surface is orientable with the same boundary as D, but may have higher genus. If D has area strictly less than $L^2/4\pi$ then the area of F can also be taken to be less than $L^2/4\pi$. If D has area equal to $L^2/4\pi$ then it spans a circle and is already embedded. In all other cases strict inequality applies for both D and F. $\square$

Remark: Almgren [1] established isoperimetric inequalities for varifolds and currents in arbitrary codimension, with optimal constants. In some range of dimension and codimension, these give upper bounds for the area of embedded spanning surfaces and spanning submanifolds, but with unspecified topological type.

References

[1] F. J. Almgren, Optimal isoperimetric inequalities, *Indiana Univ. Math. J.* **35** (1986), 451–547.

[2] F. J. Almgren and W. P. Thurston, Examples of unknotted curves that only bound surfaces of high genus within their convex hulls, *Annals of Math.* **105** (1977), 527–538.

[3] D. Avis and H. ElGindy, Triangulating point sets in space, *Disc. & Comp. Geom.* **2** (1987), 99–111.

[4] E. F. Beckenbach and T. Rado, Subharmonic functions and surfaces of negative curvature, *Trans. Amer. Math. Soc.* **35** (1933), 662–682.

[5] W. Blaschke, *Vorlesungen über Differentialgeometrie I*, 3rd Ed., Springer-Verlag: Berlin 1930.

[6] Greg Buck and Jon Simon, Thickness and crossing number of knots, *Topol. Appl.* **91** (1999), 245–257.

[7] Yu. D. Burago and V. A. Zalgaller, *Geometric Inequalities*, Springer-Verlag: Berlin 1988.

[8] Jason Cantarella, Robert Kusner and John Sullivan, On the minimum ropelength of knots and links. *Invent. Math.* **150** (2002), 257–286.

[9] T. Carleman, Sur Theorie de Minimalflächen, *Math. Zeitschrift* **9** (1921), 154–160.

[10] B. Chazelle, Convex partitions of polyhedra: a lower bound and worst case optimal algorithm, *SIAM J. Comput.* **13** (1984), 488–507.

[11] M. P. Do Carmo, *Riemannian Geometry*, Birkhäuser: Boston 1992.

[12] J. Douglas, Solution of the problem of Plateau, *Trans. Amer. Math. Soc.* **33** (1931), 263–321.

[13] H. Federer, *Geometric Measure Theory*, Springer-Verlag: Berlin 1969.

[14] A. Grey, *Tubes*, Addison-Wesley: Reading, MA 1990.

[15] R. Gulliver, Regularity of minimizing surfaces of prescribed mean curvature, *Annals of Math.* **97** (1973), 275–305.

[16] J. Hass and J. C. Lagarias, The number of Reidemeister moves needed for unknotting, *J. Amer. Math. Soc.* 14 (2001), 399-428.

[17] J. Hass and J. C. Lagarias, The minimal number of triangles needed to span a polygon embedded in $\mathbb{R}^d$, pp. 509–526 in: J. Goodman-R. Pollack Festscrift Volume (B. Aronov, J. Pach, M. Sharir, Eds.), Springer-Verlag: New York 2003.

[18] J. Hass, J. C. Lagarias and N. Pippenger, The computational complexity of knot and link problems, *J. ACM* 46 (1999), no. 2, 185–211.

[19] J. Hass, J. S. Snoeyink and W.P. Thurston, The size of spanning disks for polygonal knots, *Disc. & Comp. Geom.* **29** (2003), 1–17.

[20] M. Hirsch, *Differential Topology*, Springer-Verlag: New York 1976.

[21] A. A. Kosinski, *Differential Manifolds*, Academic Press, San Diego 1993.

[22] R. A. Litherland, J. Simon, O. Durumeric, E. Rawdon, Thickness of knots, Topology Appl. **91** (1999), 233–244.

[23] F. Morgan, *Geometric Measure Theory: A Beginner's Guide*, Academic Press: New York 1988.

[24] R. Osserman, A proof of the regularity everywhere of the classical solution to Plateau's problem, *Annals of Math.* **91** (1978) 550–569.

[25] R. Osserman, The isoperimetric inequality, *Bull. Amer. Math. Soc.* **84** (1978) 1182–1238.

[26] T. Rado, *On the Problem of Plateau.* Springer, Berlin, 1933. Reproduced by Chelsea Publishing Co., New York, NY, 1951.

[27] W. Reid, The isoperimetric inequality and associated boundary problems, *J Math. Mech.* **8** (1959) 897–905.

[28] A. Weil, Sur les surfaces à coubure négative, em C. R. Acad. Sci. Paris **182** (1926), 1069–1071. [Reprinted in: A. Weil, Oeuvres Scientifiques, Vol. I. Springer-Verlag: New York, 1979, pp. 1–2.]

Department of Mathematics
University of California
Davis, CA 95616
hass@math.ucdavis.edu

AT&T Labs-Research
Florham Park, NJ 07932-0971
jcl@research.att.com

Department of Mathematics
University of California
Davis, CA 95616
wpt@math.ucdavis.edu